

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

## The Delphion Integrated View

Get Now: ☒ PDF | [More choices...](#)

Tools: Add to Work File: [Create new Work](#)

View: [INPADOC](#) | Jump to: [Top](#) ☐ Go to: [Derwent](#)

[Email](#)



**Title:** **JP8079868A2: BONE CONDUCTION MICROPHONE OUTPUT SIGNAL REPRODUCTION DEVICE**

**Derwent Title:** Microphone output reproducing device for communication appts e.g tape recorder, mobile telephone, transceiver - regenerates pseudo-output signal based on conversion rule and operation of uniting part [\[Derwent Record\]](#)

**Country:** JP Japan  
**Kind:** A (See also: [JP3306784B2](#))

**Inventor:** NAKADAI YOSHIO;  
NISHINO YUTAKA;

**Assignee:** NIPPON TELEGR & TELEPH CORP <NTT>  
[News, Profiles, Stocks and More about this company](#)

**Published / Filed:** 1996-03-22 / 1994-09-05

**Application Number:** JP1994000211584

**IPC Code:** [H04R 1/00](#); [H04R 3/00](#);

**Priority Number:** 1994-09-05 JP1994000211584

**Abstract:** PURPOSE: To provide a bone conduction microphone output signal reproduction device in which a voice signal is corrected in the unit of phoneme so as to obtain an output voice signal with high quality.

CONSTITUTION: A voice signal is collected simultaneously from a bone conduction microphone 1 and an air conduction microphone 2 and sets of voice signal waveform patterns divided in the unit of short phoneme time are stored in a conversion rule decision device 9, and when the signal waveform pattern of a collected voice is received from the bone conduction microphone 1, a pattern closest to the signal waveform pattern stored in the conversion rule decision device 9 is selected and the signal waveform pattern related to the selected signal is obtained and the patterns are combined and the result is outputted from an output terminal 16.

COPYRIGHT: (C)1996,JPO

**INPADOC** None **Get Now:** [Family Legal Status Report](#)

**Legal Status:**  
**Family:** [Show 2 known family members](#)

**Other Abstract Info:** DERABS G96-215862 DERG96-215862



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-79868

(43) 公開日 平成8年(1996)3月22日

(51) Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
H 0 4 R	1/00	3 2 7		
	3/00	3 2 0		

審査請求 未請求 請求項の数 4 O L (全 11 頁)

(21) 出願番号 特願平6-211584

(22) 出願日 平成6年(1994)9月5日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 中▲ダイ▼ 芳夫

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

(72) 発明者 西野 豊

東京都千代田区内幸町一丁目1番6号 日

本電信電話株式会社内

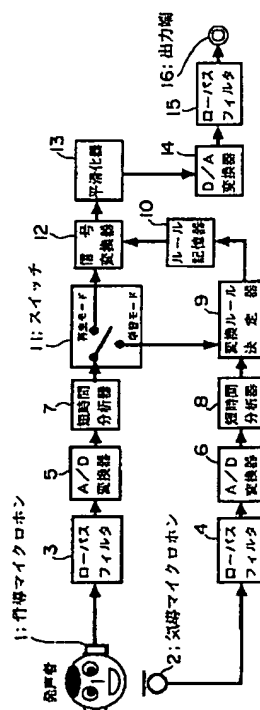
(74) 代理人 弁理士 志賀 正武

(54) 【発明の名称】 骨導マイクロホン出力信号再生装置

(57) 【要約】

【目的】 音声を音素単位で補正して高品質の出力音声を得ることができる骨導マイクロホン出力信号再生装置を提供することを目的とする。

【構成】 骨導マイクロホン1および気導マイクロホン2から同時に音声を収録し、音素レベルの短時間単位に分割した音声の信号波形パターンを組を変換ルール決定器9に記憶しておき、次に骨導マイクロホン1より収録音声の信号波形パターンが入力されたときに、変換ルール決定器9に記憶した信号波形パターンから最も近いものを選択し、選択された信号に対応付けられた信号波形パターンを導出し、それぞれを接合して出力端16より出力する。



1

## 【特許請求の範囲】

【請求項 1】 骨導マイクロホンと、骨導マイクロホン出力信号を所定の短時間毎に分割する手段と、気導マイクロホンと、気導マイクロホン出力信号を前記所定の短時間毎に分割する手段と、前記所定の短時間の骨導マイクロホン出力信号と前記所定の短時間の気導マイクロホン出力信号との対応を求め骨導マイクロホン出力信号から気導マイクロホン出力信号への前記所定の短時間単位での信号変換ルールを決定する手段と、前記信号変換ルールを記憶する手段と、該手段に記憶された信号変換ルールに基づいて前記所定の短時間の骨導マイクロホン出力信号から前記所定の短時間の疑似気導マイクロホン出力信号を生成して出力する手段と、前記所定の短時間の疑似気導マイクロホン出力信号のそれぞれを接合して長時間の疑似気導マイクロホン出力信号を得る手段とを具備し、

前記骨導マイクロホンおよび前記気導マイクロホンのそれぞれより同時収録して前記所定の短時間分割を施した各音声信号波形について 1 対 1 の対応を求め、前記所定の短時間の骨導マイクロホン出力信号から前記所定の短時間の疑似気導マイクロホン出力信号への変換ルールとして記憶し、該変換ルールおよび前記所定の短時間の骨導マイクロホン出力信号に基づいて得られる前記所定の短時間の疑似気導マイクロホン出力信号を接合して長時間の信号波形を得ることにより前記長時間の疑似気導マイクロホン出力信号を再生することを特徴とする骨導マイクロホン出力信号再生装置。

【請求項 2】 骨導マイクロホンと、骨導マイクロホン出力信号を所定の短時間毎に分割する手段と、前記所定の短時間の骨導マイクロホン出力信号を特徴抽出して基本周波数と声道特徴パラメータとを導出する手段と、気導マイクロホンと、気導マイクロホン出力信号を前記所定の短時間毎に分割する手段と、前記所定の短時間の気導マイクロホン出力信号を特徴抽出して声道特徴パラメータを導出する手段と、前記骨導マイクロホン出力信号の声道特徴パラメータと前記気導マイクロホン出力信号の声道特徴パラメータとの対応を求め、前記骨導マイクロホン出力信号の声道特徴パラメータから前記気導マイクロホン出力信号の声道特徴パラメータへの前記所定の短時間単位での変換ルールを決定する手段と、前記変換ルールを記憶する手段と、該手段に記憶された変換ルールに基づいて前記骨導マイクロホン出力信号の声道特徴パラメータから疑似気導マイクロホン出力信号の声道特徴パラメータを生成して出力する手段と、前記疑似気導マイクロホン出力信号の声道特徴パラメータと前記骨導マイクロホン出力信号の基本周波数成分とから前記所定の短時間の疑似気導マイクロホン出力信号を合成する手段と、前記所定の短時間の疑似気導マイクロホン出力信号のそれぞれを接合して長時間の疑似気導マイクロホン出力信号を得る手段とを具備し、

2

前記骨導マイクロホンと前記気導マイクロホンのそれぞれより同時収録して前記所定の短時間毎に分割した各音声信号波形について、声道特徴パラメータの抽出を施した上で 1 対 1 の対応を求め、前記骨導マイクロホン出力信号の声道特徴パラメータから前記疑似気導マイクロホン出力信号の声道特徴パラメータへの変換ルールとして記憶し、該変換ルールおよび前記骨導マイクロホン出力信号の声道特徴パラメータを用いて得たパラメータに基づいて得られる声道特徴パラメータと前記骨導マイクロホン出力信号のピッチ成分とから得られる前記所定の短時間の疑似気導マイクロホン出力信号を接合して長時間の信号波形を得ることにより前記長時間の疑似気導マイクロホン出力信号を再生することを特徴とする骨導マイクロホン出力信号再生装置。

【請求項 3】 骨導マイクロホンと、骨導マイクロホン出力信号を所定の短時間毎に分割する手段と、前記所定の短時間の骨導マイクロホン出力信号に相当する疑似気導マイクロホン出力信号を得るための信号変換ルールを記憶した手段と、該手段に記憶された信号変換ルールに基づいて前記所定の短時間の骨導マイクロホン出力信号から前記所定の短時間の疑似気導マイクロホン出力信号を生成して出力する手段と、前記所定の短時間の疑似気導マイクロホン出力信号のそれぞれを接合して長時間の疑似気導マイクロホン出力信号を得る手段とを具備し、前記変換ルールおよび前記所定の短時間の骨導マイクロホン出力信号に基づいて得られる前記所定の短時間の疑似気導マイクロホン出力信号を接合して長時間の信号波形を得ることにより前記長時間の疑似気導マイクロホン出力信号を再生することを特徴とする骨導マイクロホン出力信号再生装置。

【請求項 4】 骨導マイクロホンと、骨導マイクロホン出力信号を所定の短時間毎に分割する手段と、前記所定の短時間の骨導マイクロホン出力信号を特徴抽出して基本周波数と声道特徴パラメータとを導出する手段と、前記所定の短時間の骨導マイクロホン出力信号に相当する疑似気導マイクロホン出力信号を得るための変換ルールを記憶した手段と、前記骨導マイクロホン出力信号の声道特徴パラメータに相当する疑似気導マイクロホン出力信号の声道特徴パラメータを得るための変換ルールを記憶した手段と、該手段に記憶された変換ルールに基づいて前記骨導マイクロホン出力信号の声道特徴パラメータから疑似気導マイクロホン出力信号の声道特徴パラメータを生成して出力する手段と、前記疑似気導マイクロホン出力信号の声道特徴パラメータと前記骨導マイクロホン出力信号の基本周波数成分とから前記所定の短時間の疑似気導マイクロホン出力信号を合成する手段と、前記所定の短時間の疑似気導マイクロホン出力信号のそれぞれを接合して長時間の疑似気導マイクロホン出力信号を得る手段とを具備し、

50 前記変換ルールおよび前記所定の短時間の骨導マイクロ

3

ホン出力信号の声導特徴パラメータを用いて得られる声道特徴パラメータと前記骨導マイクロホン出力信号のピッチ成分とから得られる前記所定の短時間の疑似気導マイクロホン出力信号を接合して長時間の信号波形を得ることにより前記長時間の疑似気導マイクロホン出力信号を再生することを特徴とする骨導マイクロホン出力信号再生装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、骨導マイクロホン出力信号再生装置に関するものである。

【0002】

【従来の技術】移動電話、トランシーバ等の通信機器やテープレコーダなどの音声収録機器を使用するときに、発声者の音声を收音するために、空気中の振動を收音するマイクロホン、いわゆる気導マイクロホンが用いられることが多い。しかしながら、気導マイクロホンを用いて、工事現場などの高レベルの騒音が放射される場所で発声者の音声を収録しようとした場合、その音声には騒音が重畳し、送話に十分なS/N（音声信号対雑音比）を得ることができない。そこで、このような送話環境においては、骨導マイクロホンが使用される。

【0003】骨導マイクロホンは、骨伝導マイクロホンとも呼ばれるが、音声発生時の声帯振動によって生じる骨の振動を額、顎、頬、耳孔などで収録し、実際の音声の代用の信号として利用するための振動ピックアップのひとつである。高レベルの騒音下では騒音によって人間の骨も振動するため、これが骨導マイクロホン出力音声信号に重畳する事が観測されるが、それでも音源、すなわち声帯振動から近接した位置で音声を収録できるため、気導マイクロホンに比べて高いS/Nを得ることができ、高騒音下の音声入力手段として有効である。

【0004】ところが、骨導マイクロホンは、気導マイクロホンに比べて、周波数特性上いくつかの問題点を有している。第1の問題点は、受話信号について平坦な周波数特性が得られず、低域が強調された音声となり易い点である。第2の問題点は、声帯振動で生じる音声と、声帯以外の発生帰還を介して生じる音声とで、骨導マイクロホン出力音声信号のパワーが、気導音声の場合と異なった特徴を示す点である。例えば、耳孔に骨導マイクロホンを配置した場合、鼻腔を振動させる撥音「ん」については、鼻腔と耳孔との位置が近接しているため、声帯振動で生じる母音よりも高レベルの信号として検出され、気導音声の場合と比較して違和感のある音声となる。また、第3の問題点は、骨導マイクロホンの材質、形状や装着状態によっては、マイクロホンのユニットと皮膚との摩擦によって生じる不要音を拾いやすく、この不要音が口蓋の開閉によって常時生じる雑音として骨導マイクロホン送話音声に重畳されるという点である。

【0005】従って、骨導マイクロホンによって通常の

4

マイクロホン收音程度の明瞭な音声を得ようとした場合、骨導マイクロホン出力音声信号の各音素（または音素と同等レベルの短時間区間）ごとに周波数特性を平坦化し、音声パワーを調整し、また不要雑音を除去して再度音声合成するような信号処理技術を用いることが必要になる。このような信号処理技術として、従来より、骨導マイクロホン出力音声信号にアクティブフィルタによる補正を施して音声品質を改善する試みが行われてきた。

【0006】図6はこのフィルタ補正の例である。ここで、発声者の音声を骨導マイクロホンで収録したものを骨導音声、また音声を通常の気導マイクロホンで収録したものを気導音声と呼ぶことにする。まず、発声者の骨導音声と気導音声とをそれぞれ骨導マイクロホン1および気導マイクロホン2を使用して同時収録し、これらを一旦、テープレコーダ3などで記憶する。記憶した各々の音声波形について長時間平均スペクトルを観測し、骨導マイクロホン1での收音波形に対する気導マイクロホン2での收音波形の特性の差異を長時間スペクトル計算部4で得る。そこで、この差分特性を実現するフィルタをフィルタ部5で実現すれば、以降、骨導マイクロホン1の收音音声はフィルタ部5を通じ、気導マイクロホン2での收音音声に相当する疑似気導音声となって出力端6より得られるというものである。

【0007】

【発明が解決しようとする課題】しかしながら、上述した従来の改善方法は、長時間平均値としての音声特性の改善を行うものであって、各音節毎に正しく修正するものではない。より正確な補正を行うためには、骨導音声を音素（または音素と同等レベルの短時間）単位で分解した上で、これに予め各音素単位で求めておいた骨導音声から気導音声への音声補正フィルタ処理を施して、音声を再生成する方法が望まれる。

【0008】ところが、骨導音声には、前述したように骨導マイクロホンと皮膚との摩擦音が重畳している。したがって、前述した骨導音声を信号処理的に補正して使用する方式では、骨導音声を音素単位で分割したとしても、耳障りな雑音が残ってしまうという問題がある。本発明は上述した事情に鑑みて為されたものであり、音声を音素単位で補正して高品質の出力音声を得ることができる骨導マイクロホン出力信号再生装置を提供することを目的とする。

【0009】

【課題を解決するための手段】請求項1に記載の骨導マイクロホン出力信号再生装置は、骨導マイクロホンと、骨導マイクロホン出力信号を所定の短時間毎に分割する手段と、気導マイクロホンと、気導マイクロホン出力信号を前記所定の短時間毎に分割する手段と、前記所定の短時間の骨導マイクロホン出力信号と前記所定の短時間の気導マイクロホン出力信号との対応を求め骨導マイク

5

ロホン出力信号から気導マイクロホン出力信号への前記所定の短時間単位での信号変換ルールを決定する手段と、前記信号変換ルールを記憶する手段と、該手段に記憶された信号変換ルールに基づいて前記所定の短時間の骨導マイクロホン出力信号から前記所定の短時間の疑似気導マイクロホン出力信号を生成して出力する手段と、前記所定の短時間の疑似気導マイクロホン出力信号のそれぞれを接合して長時間の疑似気導マイクロホン出力信号を得る手段とを具備し、前記骨導マイクロホンおよび前記気導マイクロホンのそれぞれより同時収録して前記所定の短時間分割を施した各音声信号波形について1対1の対応を求め、前記所定の短時間の骨導マイクロホン出力信号から前記所定の短時間の疑似気導マイクロホン出力信号への変換ルールとして記憶し、該変換ルールおよび前記所定の短時間の骨導マイクロホン出力信号に基づいて得られる前記所定の短時間の疑似気導マイクロホン出力信号を接合して長時間の信号波形を得ることにより前記長時間の疑似気導マイクロホン出力信号を再生することを特徴としている。

【0010】請求項2に記載の骨導マイクロホン出力信号再生装置は、骨導マイクロホンと、骨導マイクロホン出力信号を所定の短時間毎に分割する手段と、前記所定の短時間の骨導マイクロホン出力信号を特徴抽出して基本周波数と声道特徴パラメータとを導出する手段と、気導マイクロホンと、気導マイクロホン出力信号を前記所定の短時間毎に分割する手段と、前記所定の短時間の気導マイクロホン出力信号を特徴抽出して声道特徴パラメータを導出する手段と、前記骨導マイクロホン出力信号の声道特徴パラメータと前記気導マイクロホン出力信号の声道特徴パラメータとの対応を求め、前記骨導マイクロホン出力信号の声道特徴パラメータから前記気導マイクロホン出力信号の声道特徴パラメータへの変換ルールを決定する手段と、前記変換ルールを記憶する手段と、該手段に記憶された変換ルールに基づいて前記骨導マイクロホン出力信号の声道特徴パラメータから疑似気導マイクロホン出力信号の声道特徴パラメータを生成して出力する手段と、前記疑似気導マイクロホン出力信号の声道特徴パラメータと前記骨導マイクロホン出力信号の基本周波数成分とから前記所定の短時間の疑似気導マイクロホン出力信号を合成する手段と、前記所定の短時間の疑似気導マイクロホン出力信号のそれぞれを接合して長時間の疑似気導マイクロホン出力信号を得る手段とを具備し、前記骨導マイクロホンと前記気導マイクロホンのそれぞれより同時収録して前記所定の短時間毎に分割した各音声信号波形について、声道特徴パラメータの抽出を施した上で1対1の対応を求め、前記骨導マイクロホン出力信号の声道特徴パラメータから前記疑似気導マイクロホン出力信号の声道特徴パラメータへの変換ルールとして記憶し、該変換ルールおよび前記骨導マイクロホン出力信号の声道特徴パラメータ

6

タを用いて得たパラメータに基づいて得られる声道特徴パラメータと前記骨導マイクロホン出力信号のピッチ成分とから得られる前記所定の短時間の疑似気導マイクロホン出力信号を接合して長時間の信号波形を得ることにより前記長時間の疑似気導マイクロホン出力信号を再生することを特徴としている。

【0011】請求項3に記載の骨導マイクロホン出力信号再生装置は、骨導マイクロホンと、骨導マイクロホン出力信号を所定の短時間毎に分割する手段と、前記所定の短時間の骨導マイクロホン出力信号に相当する疑似気導マイクロホン出力信号を得るための信号変換ルールを記憶した手段と、該手段に記憶された信号変換ルールに基づいて前記所定の短時間の骨導マイクロホン出力信号から前記所定の短時間の疑似気導マイクロホン出力信号を生成して出力する手段と、前記所定の短時間の疑似気導マイクロホン出力信号のそれぞれを接合して長時間の疑似気導マイクロホン出力信号を得る手段とを具備し、前記変換ルールおよび前記所定の短時間の骨導マイクロホン出力信号に基づいて得られる前記所定の短時間の疑似気導マイクロホン出力信号を接合して長時間の信号波形を得ることにより前記長時間の疑似気導マイクロホン出力信号を再生することを特徴としている。

【0012】請求項4に記載の骨導マイクロホン出力信号再生装置は、骨導マイクロホンと、骨導マイクロホン出力信号を所定の短時間毎に分割する手段と、前記所定の短時間の骨導マイクロホン出力信号を特徴抽出して基本周波数と声道特徴パラメータとを導出する手段と、前記所定の短時間の骨導マイクロホン出力信号に相当する疑似気導マイクロホン出力信号を得るための変換ルールを記憶した手段と、前記骨導マイクロホン出力信号の声道特徴パラメータに相当する疑似気導マイクロホン出力信号の声道特徴パラメータを得るための変換ルールを記憶した手段と、該手段に記憶された変換ルールに基づいて前記骨導マイクロホン出力信号の声道特徴パラメータから疑似気導マイクロホン出力信号の声道特徴パラメータを生成して出力する手段と、前記疑似気導マイクロホン出力信号の声道特徴パラメータと前記骨導マイクロホン出力信号の基本周波数成分とから前記所定の短時間の疑似気導マイクロホン出力信号を合成する手段と、前記所定の短時間の疑似気導マイクロホン出力信号のそれぞれを接合して長時間の疑似気導マイクロホン出力信号を得る手段とを具備し、前記変換ルールおよび前記所定の短時間の骨導マイクロホン出力信号の声道特徴パラメータを用いて得られる声道特徴パラメータと前記骨導マイクロホン出力信号のピッチ成分とから得られる前記所定の短時間の疑似気導マイクロホン出力信号を接合して長時間の信号波形を得ることにより前記長時間の疑似気導マイクロホン出力信号を再生することを特徴としている。

【0013】

【作用】請求項1に記載の骨導マイクロホン出力信号再生装置では、骨導マイクロホンを装着した発声者について、予め、骨導マイクロホンより骨導音声および気導マイクロホンより気導音声を同時収録しておき、得られた骨導音声信号および気導音声信号をそれぞれ短時間分割し、短時間の骨導音声信号と気導音声信号との対応を求め、信号変換ルールとして記憶する。次に骨導マイクロホンより骨導音声が入力されたときに、得られた骨導音声信号を短時間分割し、先に記憶した信号変換ルールによって疑似骨導音声信号へ変換し、これらを接合して長時間の疑似気導音声信号を再生する。

【0014】請求項2に記載の骨導マイクロホン出力信号再生装置では、骨導マイクロホンを装着した発声者について、予め、骨導マイクロホンより骨導音声および気導マイクロホンより気導音声を同時収録しておき、得られた骨導音声信号および気導音声信号をそれぞれ短時間分割して信号分析を行う。その結果、骨導および気導それぞれの音声信号より短時間単位での声導特徴パラメータが得られ、これら各音声信号の声導特徴パラメータ間の対応を求め、変換ルールとして記憶する。次に、骨導マイクロホンより骨導音声が入力されたときに、得られた骨導音声信号を短時間分割し信号分析して得られる声導特徴パラメータと先に記憶した変換ルールとに基づいて、骨導音声信号の疑似声導特徴パラメータを導出し、このパラメータと元の骨導音声信号の信号分析によって得られる基本周波数とを用いて、短時間の疑似骨導音声信号を合成し、これらを接合して長時間の疑似気導音声信号を再生する。

【0015】請求項3に記載の骨導マイクロホン出力信号再生装置では、骨導マイクロホンより骨導音声が入力されたときに、得られた骨導音声信号を所定の短時間毎に分割し、前記所定の短時間の骨導マイクロホン出力信号に相当する疑似気導マイクロホン出力信号を得るための普遍的な信号変換ルールによって前記所定の短時間の疑似骨導音声信号へ変換し、これらを接合して長時間の疑似気導音声信号を再生する。

【0016】請求項4に記載の骨導マイクロホン出力信号再生装置では、骨導マイクロホンより骨導音声が入力されたときに、得られた骨導音声信号を所定の短時間毎に分割して特徴抽出し、これにより得られる声道特徴パラメータと予め設定された普遍的な変換ルールとによって得られる声道特徴パラメータと、前記所定の短時間の骨導マイクロホン出力信号のピッチ成分とを合成して前記所定の短時間の疑似骨導音声信号を生成し、これらを接合して長時間の疑似気導音声信号を再生する。

【0017】

【実施例】以下、図面を参照して本発明の実施例について説明する。図1は本発明の第1の実施例による骨導マイクロホン出力信号再生装置の概略構成を示すブロック図である。図1において、1は骨導マイクロホンであ

り、顔の部位、例えば、額、顎、頬、耳孔などに装着され、骨や皮膚に伝達される発声者の声帯振動を収録するものである。2は気導マイクロホンであり、空気伝搬する発声者の肉声信号を収録するものであり、すなわち一般的なマイクロホンである。

【0018】3および4はローパスフィルタであり、それぞれ骨導マイクロホン1および気導マイクロホン2の出力信号に対してエリアシング歪みを防止するためのものである。ローパスフィルタ3、4のカットオフ周波数は、最終的に得ようとする疑似気導音声の周波数帯域を元の骨導音声と同一帯域にしようとするものであれば、それぞれ同一の値、例えば、4kHzである。また、最終的に得ようとする疑似気導音声の周波数帯域を元の骨導音声の帯域よりも拡大しようとする場合には、カットオフ周波数は、例えば、ローパスフィルタ3については4kHz、ローパスフィルタ4については7kHzというように、それぞれ異なった値としてもよい。

【0019】5および6はA/D変換器であり、それぞれローパスフィルタ3および4の出力について、後段で行われる信号処理を容易にするためにA/D変換を施すものである。各A/D変換器5、6は、それぞれのサンプリング周波数の音声の特徴が明確に現れる周波数帯域を含み、かつ、ローパスフィルタ3、4のカットオフ周波数に対してナイキストの標準化定理を満たす関係であれば良い。また、A/D変換器5、6の量子化ビット数は、音声の特徴が明確に表れ、量子化歪が少ないものであれば良い。

【0020】すなわち例えば、ローパスフィルタ3、4のそれぞれのカットオフ周波数を4kHz同一とした場合のA/D変換器5および6のサンプリング周波数および量子化ビット数は、例えば、8kHzサンプリング、12ビット線形量子化で同一となる。また、ローパスフィルタ3については4kHz、ローパスフィルタ4については7kHzのように、各ローパスフィルタ3、4のカットオフ周波数が異なっている場合、A/D変換器5については、例えば8kHzサンプリング、12ビット線形量子化となり、A/D変換器6については、例えば16kHzサンプリング、16ビット線形量子化となる。

【0021】7および8は短時間分析器であり、それぞれA/D変換器5、6より得られる骨導音声信号および気導音声信号を短時間区間単位に分割する。この分割単位は、各短時間分析器7、8で同一の値をとり、音素あるいは音韻レベルの時間長、例えば32msecとする。また例えば、後述する平滑化器13において窓関数を乗じるために信号パワーの損失が生じるような場合には、短時間分析器7、8では、波形の一部を重複させながら分割することによって窓関数での損失を避けるような処理を行う。この処理により分割される波形の例を図2に示す。この図に示す例では、原音声波形の一部を重

複させながら分割波形パターンA、B、Cを生成している。

【0022】再び図1において、9は変換ルール決定器であり、短時間分析器7、8で得られた短時間の骨導音声信号と気導音声信号との対応を学習して記憶するものである。すなわち短時間の骨導音声信号を $a(n)$ 、 $a(n)$ と同時に収録した短時間の気導音声信号を $b(n)$ とすると、変換ルール決定器9は、 $a(n)$ と $b(n)$ との組を信号変換ルールとして決定する。なお、 $n$ は記憶するルールの番号を示しており、ここではルールの個数は最大で1000となるものとする。

【0023】10はルール記憶器であり、変換ルール決定器9で決定された変換ルールを記憶し、後段の信号変換器12に与えるものである。11は短時間分析器7からの骨導音声信号の出力先を切り替えるスイッチであり、このスイッチ11により、信号変換ルールを学習する学習モードと、信号変換ルールに基づいて骨導音声信号の信号変換を行う再生モードとが切り替えられる。

【0024】12は信号変換器であり、ルール記憶器10で記憶された信号変換ルールに基づいて、短時間分析器7から出力される短時間の骨導音声信号から短時間の疑似気導音声信号を得るものである。平滑化器13は、信号変換器12の出力である短時間の疑似気導音声信号を、元の骨導音声信号の時間軸に合わせて接合し、また接合端で信号が不連続になることによって信号歪みが出ることをないように平滑化処理を施すものである。平滑化の手法としては、例えば、ハミング窓関数によって信号接合部の振幅値を0に近似した値とするものである。

【0025】14はD/A変換器であり、平滑化器13から出力されるデジタル信号を、アナログ信号に変換するものである。15はローパスフィルタであり、D/A変換器14の出力信号について、エリasing歪みを防止する。ここで、D/A変換器14はA/D変換器6と同一のサンプリング周波数及び量子化ビット数を有し、また、ローパスフィルタ15はローパスフィルタ4と同一のカットオフ周波数を有するものとする。16は最終的に疑似気導音声信号を出力する出力端である。

【0026】上述した構成による装置の動作について、学習モードと、再生モードとに分けて説明する。学習モードは、骨導音声と気導音声との対応を求めて信号変換ルールを決定するモードであり、再生モードは、信号変換ルールに基づいて、骨導音声から疑似気導音声出力を得るモードである。

#### 【0027】(1) 学習モードの動作

学習モードにおいては、スイッチ11は学習モードの方へ接続されている。このような状態において、まず、発声者が、音声信号としてあらゆる特徴が表出した語彙や文章、例えば、文献、板橋著、「音声認識用共通音声データ」、日本音響学会シンポジウム「試験用音声の標準化」予稿集、1985年、に述べられているような10

0個の日本都市名などを発声する。ここで発声者の使用する環境は、音声の特徴抽出に悪影響を与えない周囲騒音レベルの少ない室内であることが必要である。

【0028】発声された音声は、骨導マイクロホン1および気導マイクロホン2にそれぞれ同時に入力され、ローパスフィルタ3、4およびA/D変換器5、6を通じてデジタル形式の波形データに変換される。このデジタル形式の波形データは短時間分析器7、8において前述したように短時間単位で分割され、変換ルール決定器9へ送出される。変換ルール決定器9では前述したように、短時間の骨導音声信号 $a(n)$ と短時間の気導音声信号 $b(n)$ とを組み合わせて $a(n)$ から $b(n)$ への変換ルールとする。

【0029】なお、 $a(n)$ については、多数の骨導音声信号を観測した場合、類似した信号パターンが観測されるが、類似した信号パターンについては同一の $a(n)$ として扱われる。すなわち、既に変換ルール決定器9に記憶された $a(n)$ の、例えばLPC係数などのスペクトル上の特徴量を $A(n)$ とすると、新たに入力された短時間の骨導音声信号 $a(n')$ についてそのスペクトルを $A(n')$ としたとき、スペクトル上の特徴量の距離の絶対値 $|A(n) - A(n')|$ が所定のしきい値THよりも小さい場合に、この $a(n')$ は $a(n)$ と同一の骨導音声信号パターンとして分類される。このようにして、変換ルール決定器9では、ある一定個数の変換ルールが決定される。こうして得られた変換ルールはルール記憶器10に記憶され、十分な数の変換ルールが得られると学習モードが終了する。

#### 【0030】(2) 再生モードの場合

再生モードは、学習モードが終了した後に使用されるモードである。再生モードにおいては、スイッチ11は再生モードの方へ接続されており、短時間分析器7と信号変換器12とが接続されている。また、このモードでは、気導マイクロホン2、ローパスフィルタ4、A/D変換器6、短時間分析器8、変換ルール決定器9は使用されない。

【0031】再生モードでは、発声者の音声は、骨導マイクロホン1、ローパスフィルタ3、A/D変換器5を通じてデジタル形式の波形データに変換され、短時間分析器7で短時間単位に分割された後で信号変換器12へ送出される。ここで、信号変換器12に送出される骨導音声信号を $x$ とする。次に信号変換器12では、入力された $x$ とルール記憶器10で記憶した各 $a(n)$ とのスペクトル上の特徴量の距離の絶対値 $D(n)$ を求める。なお、 $D(n)$ は、 $x$ 、 $a(n)$ のスペクトル上の特徴量をそれぞれ $X$ 、 $A(n)$ とすると、 $D(n) = |X - A(n)|$ である。

【0032】ここで、 $D(n)$ が最小値となる場合の $a(n)$ が、入力信号 $x$ の疑似骨導音声信号とされ、疑似

気導音声信号  $b(n)$  が導出される。導出された  $b(n)$  は平滑化器 13 へ送出され、ここで短時間分割信号より長時間の信号へ変換される。この信号はデジタル形式の波形データであるため、D/A変換器 14 およびローパスフィルタ 15 を介してアナログ信号波形に変換され、出力端 16 より元のアナログ信号として出力される。

【0033】次に、図3は本発明の第2の実施例による骨導マイクロホン出力信号再生装置の概略構成を示すブロック図である。図3において、図1と共通する部分には同一の符号を付し、その説明を省略する。図3において、16、17はLPC分析器であり、それぞれ短時間分析器7、8の出力について線形予測分析(LPC)を行い、入力音声のピッチ周波数と、声道特徴を示すパラメータ、例えばLPC係数などと分離するものである。

【0034】ここで分離されたもののうち、骨導音声のピッチ周波数は後述するLPC合成器19へ送出され、また骨導音声および気導音声の特徴パラメータがそれぞれスイッチ11および係数変換ルール決定器9'へ送出される。係数変換ルール決定器9'は、骨導音声の特徴パラメータから気導音声の特徴パラメータへの変換ルールを決定し、ルール記憶器10へ供給するものである。

【0035】12'は係数変換器であり、係数変換ルール決定器9'で決定され、ルール記憶器10に記憶された変換ルールに基づいて、スイッチ11を介して入力される骨導音声の特徴パラメータより疑似気導音声の特徴パラメータを導出するものである。LPC合成器19は、LPC分析器17より出力された骨導音声のピッチ周波数と係数変換器12'より出力された疑似気導音声の特徴パラメータとにより、線形予測分析(LPC)合成を行って短時間の疑似気導音声の信号波形を生成するものである。

【0036】次に、図3に示す構成の骨導マイクロホン出力信号再生装置の動作について、第1の実施例と同様に、学習モードと再生モードとに分けて説明する。

### (3) 学習モードの場合

学習モードにおいては、スイッチ11は学習モードの方へ接続されている。このような状態で、まず、発声者が、第1の実施例の場合と同様に音声信号としてあらゆる特徴が表出した語彙や文章を発声する。この音声は、骨導マイクロホン1および気導マイクロホン2にそれぞれ同時に入力され、ローパスフィルタ3、4およびA/D変換器5、6を通じてデジタル形式の波形データに変換され、短時間分析器7および8において短時間単位で分割され、係数変換ルール決定器9'へ送出される。

【0037】係数変換ルール決定器9'では、まず多数に分割した骨導音声について、LPC係数などの音声特徴パラメータ  $x(t)$  (ただし  $t$  は入力時刻) を抽出して記憶し、この中から音声の特徴を広く網羅する一定個

数の代表的なパラメータ  $p(n)$  を導出する。この方法はベクトル量子化としてよく知られているものであり、具体的手法としては、例えばLBGアルゴリズムや、K-平均クラスタリング等の名称で知られているものが使用される。なお、ここでは、最終的に分類されたパラメータの個数、すなわちコードブック数  $n$  を例えば、256個とする。

【0038】次に、前出の骨導音声特徴パラメータ  $x(t)$  をこの256個の代表的パラメータ  $p(n)$  のいずれかに置換する。すなわち、 $x(t)$  と  $p(n)$  の、例えばLPCケプストラムなどのスペクトル上の特徴量をそれぞれ  $X(t)$ 、 $P(n)$  としたとき、そのスペクトル上の特徴量の距離の絶対値  $D(n) = |X(t) - P(n)|$  について、 $D(n)$  が最小値を取るときの  $p(n)$  によって  $x(t)$  を置換する。

【0039】ここで、骨導音声特徴パラメータ  $x$  と同時に収録した気導音声特徴パラメータを  $y(t)$ 、また  $x$  が  $p(n)$  に置換されたときの  $y(t)$  を  $y(t, n)$  としたとき、全ての  $t$  に対して、 $y(t, n)$  は  $p(n)$  毎に、すなわちここでは256種類に分類されるが、分類された  $y(t, n)$  について集計され、その相加平均をとって平均値  $q(n)$  が算出される。上述した操作によって、骨導音声特徴パラメータ  $p(n)$  に対する疑似気導音声特徴パラメータ  $q(n)$  の変換ルールが導出される。係数変換ルール決定器9'では、この  $p(n)$  と  $q(n)$  との組をルール記憶器10へ送出し、ルール記憶器10で記憶させる。

### 【0040】(4) 再生モードの場合

再生モードにおいては、スイッチ11は再生モードの方へ接続され、LPC分析器17と係数変換器12'とが接続される。このような状態において、発声者の音声は骨導マイクロホン1、ローパスフィルタ3、A/D変換器5を通じてデジタル形式の波形データに変換され、短時間分析器7で短時間単位に分割され、LPC分析器17によりピッチ周波数データ  $v$  と骨導音声特徴パラメータ  $x$  とに分離される。

【0041】骨導音声特徴パラメータ  $x$  については信号変換器12へ送られ、予め係数変換ルール決定器9'で算出されルール記憶器10に記憶された代表的パラメータ  $p(n)$  によって置換される。すなわち、 $x$  および  $p(n)$  のスペクトルをそれぞれ  $X$ 、 $P(n)$  としたとき、そのスペクトル距離の絶対値  $D(n) = |X - P(n)|$  について、 $D(n)$  が最小値を取るときの  $p(n)$  によって  $x$  が置換される。

【0042】ここで、ルール記憶器10で記憶したルールに基づき、 $p(n)$  から疑似気導音声特徴パラメータ  $q(n)$  が導出され、LPC合成器19へ送出される。LPC合成器19では、 $q(n)$  とLPC分析器17から出力されたピッチ周波数データ  $v$  とに基づいて、短時間単位の疑似気導音声の信号波形が生成される。生成さ

13

れた疑似気導音声信号波形は平滑化器 13 へ送出され、短時間分割信号から長時間の信号へ変換される。平滑化器 13 の出力信号はデジタル形式の波形データであるため、D/A 変換器 14 およびローパスフィルタ 15 を経由してアナログ波形に変換され、出力端 16 より元のアナログ信号として出力される。

【0043】ここで、上述した第 1 および第 2 の実施例のそれぞれのルール記憶器 10 で記憶した各変換ルールが、いかなる発声者に対しても普遍的な変換結果をもたらすルールであれば、気導マイクロホンと、気導マイクロホンで収録した気導音声信号を信号処理する部分と、変換ルールを算出して決定する部分とが必要になり、骨導マイクロホン出力信号再生装置の構成はより簡易になる。このような構成の骨導マイクロホン出力信号再生装置について以下に説明する。

【0044】図 4 は本発明の第 3 の実施例による骨導マイクロホン出力信号再生装置の概略構成を示すブロック図であり、図 1 と共通する部分には同一の符号を付し、その説明を省略する。この図に示す装置は、図 1 に示すものから、気導マイクロホン 2、ローパスフィルタ 4、A/D 変換器 6、短時間分析器 8、変換ルール決定器 9、およびスイッチ 11 を取り去った構成となっている。ただし、図 4 のルール記憶部 10 には、予め、いかなる発声者に対しても普遍的な変換結果をもたらす変換ルールが記憶されている。

【0045】また、図 5 は本発明の第 4 の実施例による骨導マイクロホン出力信号再生装置の概略構成を示すブロック図であり、図 3 と共通する部分には同一の符号を付し、その説明を省略する。この図に示す装置は、図 3 に示すものから、気導マイクロホン 2、ローパスフィルタ 4、A/D 変換器 6、短時間分析器 8、LPC 分析器 18、係数変換ルール決定器 9、およびスイッチ 11 を取り去った構成となっている。ただし、図 5 のルール記憶部 10 には、予め、いかなる発声者に対しても普遍的な変換結果をもたらす変換ルールが記憶されている。

【0046】上述した第 3 および第 4 の実施例による骨導マイクロホン出力信号再生装置では、予めルール記憶部 10 に変換ルールが記憶されているため、第 1 および第 2 の実施例における学習モードが存在しない。したがって、第 1 および第 2 の実施例における再生モードと同様の動作のみが行われる。

【0047】以上説明したように、図 1、図 3、図 4、および図 5 に示す構成によって、骨導マイクロホン 1 で収録した音声に対し、予め作成した変換ルールに基づいて、音素レベルの短時間単位で骨導音声から疑似気導音声への変換処理が行われる。したがって、骨導音声に含まれる雑音が疑似気導音声へ与える影響を除去することができ、従来の時間平均値に基づく一定のフィルタ特性での補正方式に比べて、優れた特性の音声を疑似気導音声として得ることができる。

14

【0048】さらに、図 1、図 3 に示す構成では、骨導マイクロホン 1 とマイクロホン 2 とで同時に音声を収録できるため、それぞれのマイクロホンで収録した信号について 1 体 1 対応をとることができる。この対応は、発声者に応じて求めることができるため、骨導音声から疑似気導音声への変換処理を極めて高い確度で行うことができる。

【0049】また、図 3 および図 5 に示す構成では、骨導音声および気導音声をそれぞれ各種の信号分析技術によって基本周波数と声導特徴パラメータとに分離し、声導特徴パラメータの使用によって変換ルールを生成するように構成されている。したがって、ルール記憶部 10 の記憶容量を低減し、また音素単位で短時間分割した音声を接合する際に音声品質を良好とすることができる。さらに、図 4 および図 5 に示す構成では、予め不特定多数の発声者について普遍的な変換ルールを作成しておくようにしたため、装置構成を簡素とすることができるとともに、装置の使用者も変換ルールを作成（学習）するための手間を省くことができる。

【0050】

【発明の効果】本発明では、従来の骨導マイクロホンでは収音できなかった高い周波数の信号成分についても正確に再生できるという効果がある。また、短時間（音素）単位での変換を行うことにより、音声の平均スペクトルの差分により補正していた従来の方法に比べて、音素毎に最適な音声を再生できるという効果がある。さらに、骨導マイクロホンで収録した音声を補正用の音声の現信号として使用しないため、骨導マイクロホンで収音した音声に重畳する不要雑音が変換後の音声に残留しない。すなわち、出力信号から不要雑音の影響を除去することができるという効果がある。また、予め発声者に普遍的な信号変換ルールを記憶しておくことにより、使用者毎の学習操作を不要とすることができるという効果がある。

【図面の簡単な説明】

【図 1】本発明の第 1 の実施例による骨導マイクロホン出力信号再生装置の概略構成を示すブロック図である。

【図 2】音声信号の短時間分析を説明するための図である。

【図 3】本発明の第 2 の実施例による骨導マイクロホン出力信号再生装置の概略構成を示すブロック図である。

【図 4】本発明の第 3 の実施例による骨導マイクロホン出力信号再生装置の概略構成を示すブロック図である。

【図 5】本発明の第 4 の実施例による骨導マイクロホン出力信号再生装置の概略構成を示すブロック図である。

【図 6】従来の骨導マイクロホン出力信号再生装置を説明するためのブロック図である。

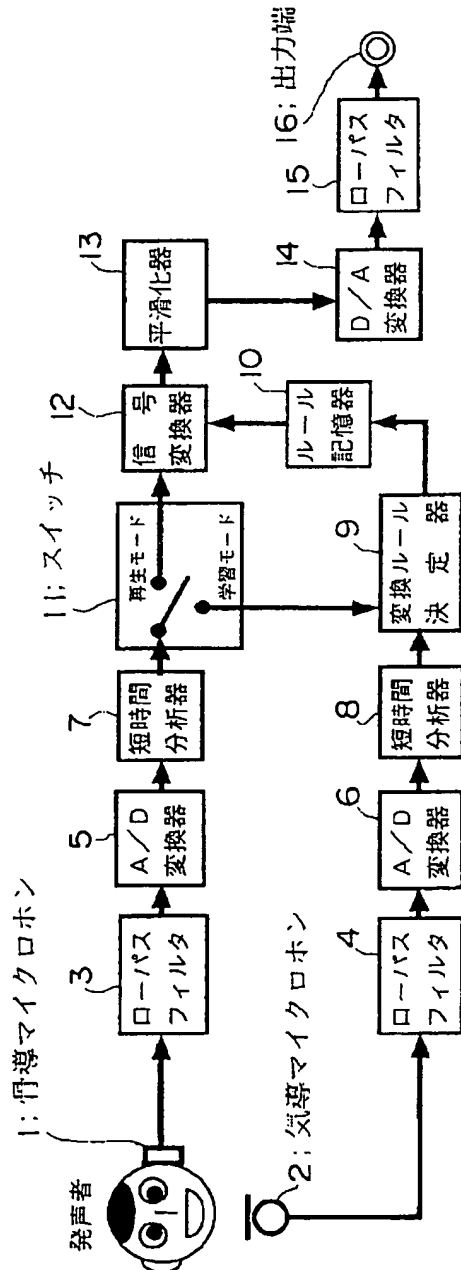
【符号の説明】

1…骨導マイクロホン、2…気導マイクロホン、3、4、15…ローパスフィルタ、5、6…A/D 変換器、

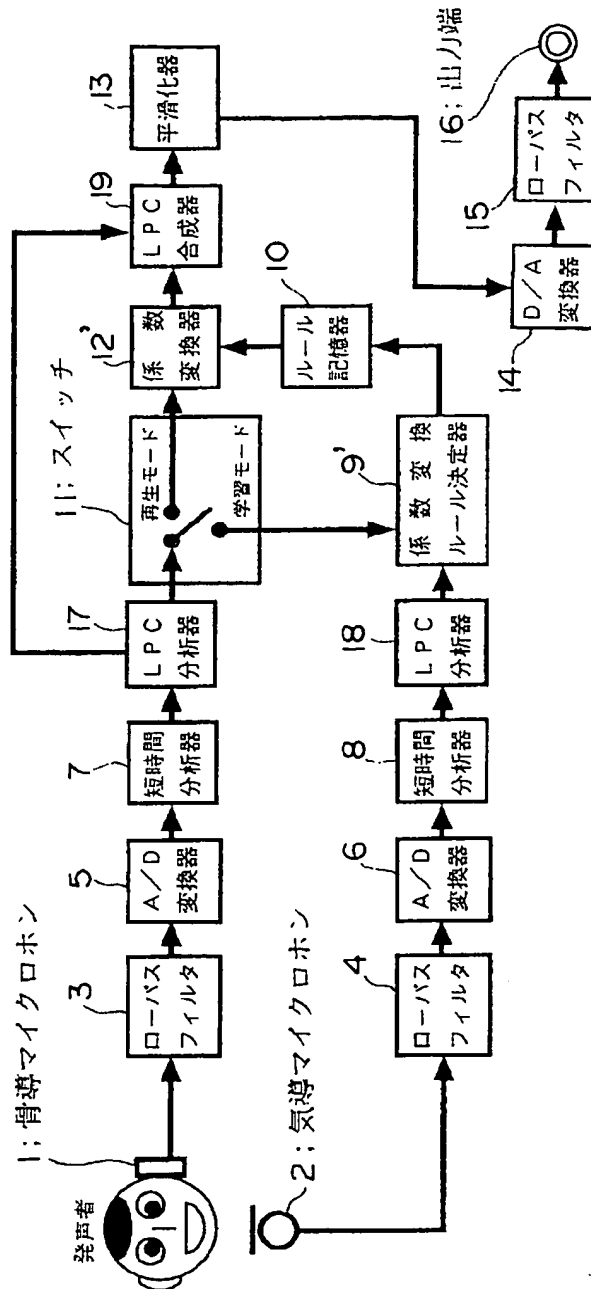
7, 8...短時間分析器、9...変換ルール決定器、9'...  
係数変換ルール決定器、10...ルール記憶器、11...ス  
イッチ、12...信号変換器、12'...係数変換器、13

...平滑化器、14...D/A変換器、16...出力端、1  
7, 18...LPC分析器、19...LPC合成器。

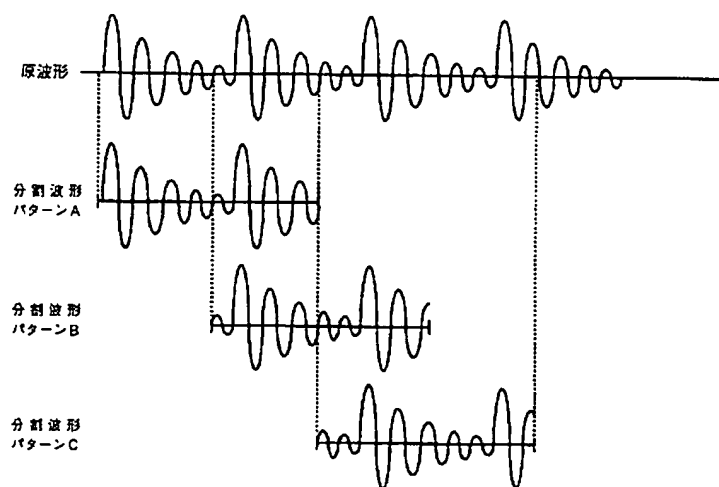
【図1】



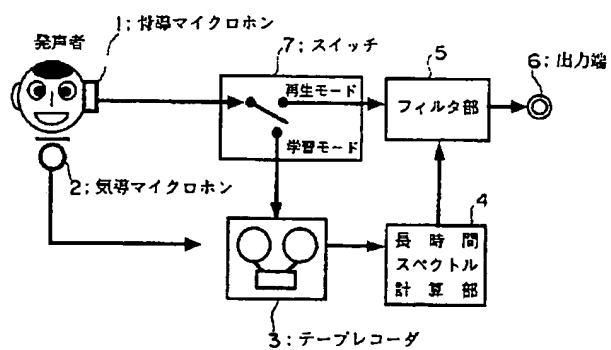
【図3】



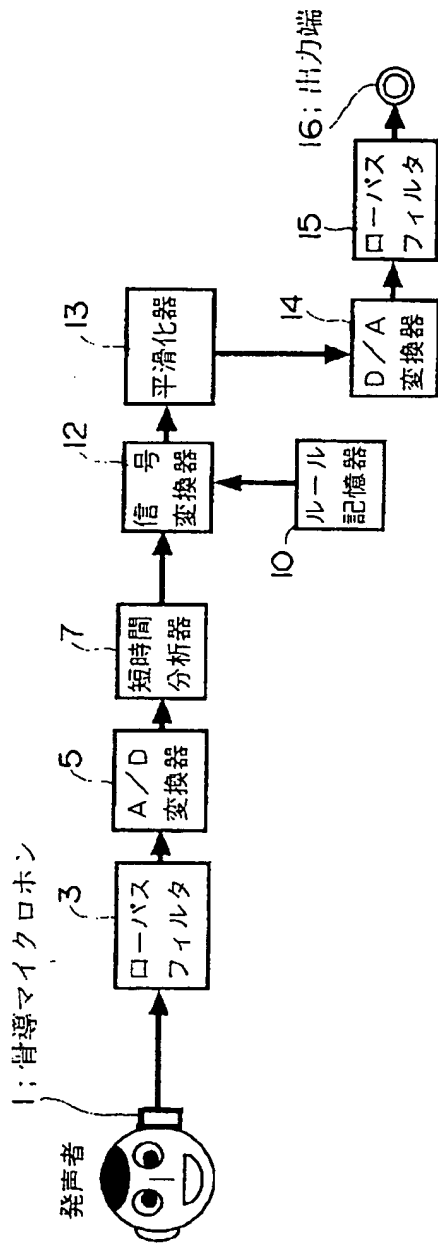
【図2】



【図6】



【図4】



【図5】

